BIOINFORMATION Discovery at the interface of physical and biological sciences

open access

www.bioinformation.net

Hypothesis

Volume 8(9)

A Phylogenetic analysis of Heparanase (HPSE) gene

Abjal Pasha Shaik¹^{\$*}, Abbas H Alsaeed¹^{\$} & Asma Sultana²^{\$}

¹Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Kingdom of Saudi Arabia; ²Janaki Nagar Colony, No: 45, Tolichowki, Hyderabad, India; Abjal Pasha Shaik – Email: afzalshaik@gmail.com; Phone: +966544023766; Fax: +96614684995; *Corresponding author \$-All authors contribute equally

Received April 21, 2012; Accepted April 27, 2012; Published May 15, 2012

Abstract:

The Current Study aimed to investigate the possible role of Heparanase protein (HPSE-1, [Entrez Pubmed ref|NP_001092010.1], heparanase isoform 1 preproprotein [Homo sapiens]) in evolution by studying the phylogenetic relationship and divergence of HPSE-1 gene using computational methods. The Human HPSE protein sequences from various species were retrieved from GenBank database and were compared using sequence alignment. Multiple sequence alignment was done using Clustal-W with defaults and phylogenetic trees for the gene were built using neighbor-joining method as in BLAST 2.2.26+ version. A total of 112 BLAST hits were found for the heparanase query sequence and these hits showed putative conserved domain, Glyco_hydro_79n superfamily. We then narrowed down the search by manually deleting the proteins which were not HPSE-1. These sequences were then subjected to phylogenetic analyses using the PhyML and TreeDyn software. Our study indicated that HPSE-1 is a conserved protein in classes Mammalia, Aves, Amphibia, Actinopterygii and Insecta emphasizing its importance in the physiology of cell membranes. Occurrence of this gene in evolution with conserved sites strengthens the role of HPSE-1 gene and helps in better understanding the biochemical processes that may lead to cancer.

Key words: HPSE-1, Heparanase, Protein sequences, Phylogeny, Biological role, Sequence alignment

Background:

Cancer remains an important cause of chronic illness. Better understanding of various genes and proteins and their implications at the cellular and molecular levels helps in identifying appropriate preventive/diagnostic measures. Studies in our lab on heparanase (HPSE-1 [Entrez Pubmed ref |NP_001092010.1 |, heparanase isoform 1 preproprotein [Homo sapiens]) gene polymorphisms associated with cancer have provided a preliminary understanding of the role of this gene in the pathophysiology of cancer (unpublished). Heparanase is a predominant endoglycosidase that cleaves heparan sulfate which forms a principal polysaccharide in the basement membrane and extracellular matrix. Cleavage of heparan sulfate disintegrates the structural integrity of the basement membrane and releases angiogenic and growthpromoting mediators [1, 2]. Heparanase is synthesized as an inactive 65 kDa proheparanase in the golgi apparatus and later transferred to endosomes/lysosomes for transport to the cellsurface [3]. The human heparanase gene (HPSE) is located on chromosome 4 (4q21.3) and is expressed as a 5- and 1.7-kb mRNA by alternative splicing [4]. The heparanase mRNA encodes a 61.2-kDa protein with 543 amino acids. This proenzyme is then post-translationally cleaved into 8 (Gln 36 to Glu 109) - and 50 (Lys 158 to Ile 543)-kDa subunits that noncovalently associate to form the active heparanase [5, 6]. The formation of this heterodimer is essential for heparanase activity. Penetration of the endothelial layer of cell in the surface of blood vessels is an important process in tumor metastases . SHA proteoglycans are important components of this layer [7]. Therefore, increased metastatic potential may correspond to increased heparanase activity [8]. Because of this inherent role of heparanase, the inhibition of its activity can be a potential target for anti-cancer therapies. Heparanase also elicits an indirect neovascular response by releasing extracellular matrix resident HS-bound angiogenic factors [9]. Considering the importance of HPSE-1 gene in cancer, we aimed to elucidate

if variance among this gene exists in various species during evolution by using a phylogenetic analysis of published protein sequences of these genes.

Methodology:

Data Set, Sequence Alignment and Phylogenetic Tree Building The GenBank database [10] was queried to retrieve all available protein sequences of the heparanse gene. These sequences were retrieved and saved in FASTA sequence format. These sequences were then aligned using Clustal W [11] algorithm using default parameters. The initial first-pass phylogenetic tree was constructed using Neighbour Joining [12] method (maximum sequence difference of 0.85) using Basic Local Alignment Search Tool [BLAST] pairwise alignments between a query and the database sequences searched [13]. Evolutionary distance between two sequences modeled as expected fraction of amino acid substitutions per site given the fraction of mismatched amino acids in the aligned region was taken by the software using Grishin computation [14]. Using the results from BLAST [15] we created a first-pass phylogenetic tree after which we used a purpose-built computational phylogenetic method using Phylogeny.fr software. Sequences were aligned with MUSCLE (v3.7) [16] configured for highest accuracy (MUSCLE with default settings). After alignment, ambiguous regions (i.e. containing gaps and/or poorly aligned) were removed with Gblocks (v0.91b) [17-19]. The phylogenetic tree was reconstructed using the maximum likelihood method implemented in the PhyML program (v3.0 aLRT) [15-20]. The Jones-Taylor-Thornton (JTT) substitution model was selected assuming an estimated proportion of invariant sites (of 0.108) and 4 gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data (gamma=1.163). Reliability for internal branch was assessed using the aLRT test (SH-Like). The graphical representation and edition of the phylogenetic tree were performed with TreeDyn (v198.3) [21].

1 75 150 225 300 375 450 525 543 <u>Glyco_hydro_79n superfamily</u>

Figure 1: Putative sequence of heparanase in super families. The actual alignment was detected with superfamily member pfam03662: Cd Length: 320 Bit Score: 65.50 E-value: 9.29e-12

Results:

From the NCBI GenBank database, 112 sequences of heparanse 1 isoform covering the putative conserved Glyco_hydro_79n domain were obtained and used for the construction of a firstpass phylogenetic tree. However, some of the repetitive and non sequences which were not related to HPSE-1 were deleted. The short-listed sequences majorly belonged to the classes Mammalia, Aves, Amphibia, Actinopterygii and Insecta (lancelets, bony fishes, frogs and toads, marsupials, rodents, primates, even-toed ungulates, carnivores, rabbits and hares, odd-toed ungulates, lizards, birds and placentals). The accession information for these sequences is available in Table 1 (see supplementary material). Analysis of the sequences revealed that there is a high degree of sequence similarity of heparanase protein in many of the Mammalian species used for the phylogeny reconstruction, thus implying that this protein might be more or less conserved in most of the Mammalian species. The two species belonging to Class Insecta also seem to share a high degree of sequence similarity with each other, but show very less degree of sequence similarity with the other classes. Putative conserved domains were observed in many taxa at the Glycosyl hydrolase family 79, N-terminal domain; Family of endo-beta-N-glucuronidase (Figure 1) [22]. The actual alignment was detected with superfamily member pfam03662 (E-value: 9.29e-12). BLAST produced 112 hits (Figure 2); these sequences were screened manually and only those related to the sequence in question (HPSE-1) from different taxa were retained for further analyses. This produced a total of 31 sequences from various taxa. Multiple sequence alignment results of these short-listed sequences are presented in (Figure 3). Using the PhyML program a tree was constructed for these 31 sequences, the results of which are presented in (Figure 4).

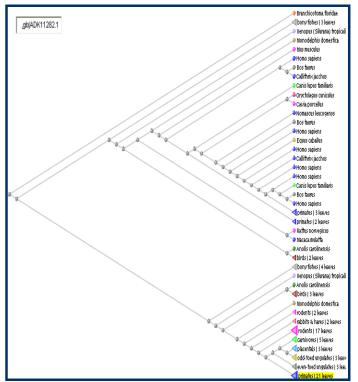


Figure 2: First pass phylogenetic tree constructed by multiple alignment using BLAST pair wise alignments: Results presented using Taxonomic name [112 hits]

Discussion:

Heparan sulfate proteoglycans (HSPGs) play a key role in the self-assembly, insolubility and barrier properties of basement membranes and extracellular matrices [23]. Hence, cleavage of heparan sulfate (HS) affects the integrity and functional state of tissues and thereby fundamental normal and pathological phenomena involving cell migration and response to changes in the extracellular micro-environment [1, 2]. Heparanase, an enzyme that degrades heparan sulfate at specific intra-chain sites is synthesized as a latent approximately 65 kDa protein that is processed at the N-terminus into a highly active approximately 50 kDa form. Experimental evidence suggests that heparanase may facilitate both tumour cell invasion and neovascularization, both critical steps in cancer progression. The enzyme is also involved in cell migration associated with inflammation and autoimmunity. Our preliminary studies (unpublished) have indicated that the HPSE gene might have association in the pathophysiology of cancer. Phylogenetics was performed in this study to evaluate the evolutionary relatedness

of heparanase protein among various groups of organisms (e.g., species, taxa, phyla). We aimed to unearth the relationship of this protein synthesized by HPSE-1 gene using computational phylogenetics to identify the seemingly similar genes in different organisms. This was done through the use of algorithms, methods and programs through phylogenetic analyses to assemble a phylogenetic tree from a set of genes, species, or other taxa [24, 25]. Neighbor-joining method was used to calculate genetic distance from multiple sequence alignments, and ClustalW was used to create trees based on distance. JTT matrix is an efficient method for generating mutation data matrices from protein sequences by using the peptide-based sequence comparison algorithm, the set sequences are clustered at the 85% identity level. The closest relating pairs of sequences are aligned, and observed amino acid exchanges tallied in a matrix [26]. Later, a phylogenetic tree is constructed from closely related sequences [20]. Several studies and especially the BAliBASE benchmark showed that MUSCLE achieved the highest ranking of any method at the time of publication. Gblocks eliminates poorly aligned positions and divergent regions (removes alignment noise). PhyML was shown to be at least as accurate as other existing phylogeny programs using simulated data, while being one order of magnitude faster. Furthermore, the system offers the possibility to control results of each step before launching the next program, so that users can modify and properly adjust parameters for a given task. This is possible in checking the "Step by step" option.

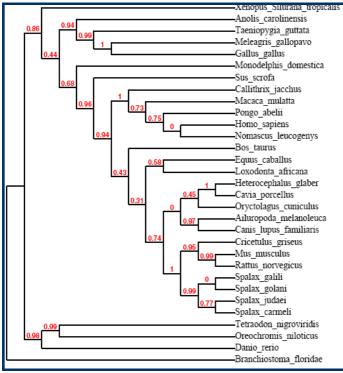
Telatil	ig pairs of sequences	arc	anglieu, and observed	ann	110				
Branchiost									
Danio reri Tettandon		Rus muscul	STAPTIONILLETT-RELEXISTERIAL TIDALATURE LITLOPELALARUSE STAPTIONILLETT-RELEXISTERIAL TIDALATURE LITLOPELALARUSE	Canis_lups		Bos_teurus	YEPOTOGENINTYIMANTIAAGEETIISYTMIHTYYMHTATMEIFTAPOTLIITYISY	Ucepheria Elpina_calia	
Cesochroni		Rattus nor Spalax jud	STATEOVILLETT-ERCEQUESESTATIACTICALATIVESCUTICALARCLES **ACAEOVILLETT-CRELICUSEDFLATTICALATIVEFCTFLOREALARCLES			Dos scrofs Crimetulus	YERGISQRABHARCKISTUTOCKY IDENTWIKYYLIKOCATHEIFLBOWETYJICA VEDWODROFYWLIKSTUKOCKY IDAL TWENYLIKO ZATWENYLIKOVUTYJICA	Allierpola	BI-90201201900
Rendpus 31		Spalan gal	ACREDVVELEPET-ORPLACVEPEPLECTIONALATOPPELTFLOEPHCAALASCLEP	Branchiget	ILLNERWYTTAAGIAPWYNCOGTAEIPLIPEPTEESTTRIAGPELSICALIROOTK	Mus muscul	YEPOTOGPHORTWOLLASTLANGER TOOLTHHEYYLNER LATHOPLOSOVLDTFTLAV	Carcia, Juga	
Realis_car Taeniopygi		Spalar_ptl	++AQAEDWELEPTT-QRPLIEUSBEFLETTIDAKLADDPRFLTFLOEDKLAALANDLEP	Datio_reri		Battus tor	ADAD OĞĞĞÜRANITYRUTARODILIDIT DIRRALINIDAL MUDALITIRA	The State Con-	
Meleagris		fpalas_car Callithris	 ACAEDWELEFST-ORFLRUSPISTISTISMILATIERFLTFLÖSPKLAALAKGLEF GFACAEDAVELDFTT-OCFLRUSPISTISTISMILATIERFLUTLGEPKLATIAALLEF 	Tetraodon Opeochromi	LITEROCCORE-	Spalax jud Spalax pal	YERVOLPRORTYKLASSERADDELDSYTNEHYTLAGRATIKDELSPULDTYLLDV YERVOLPRORTYKLASSERADDELDSYTNEHYTLAGRATIKDELSPULDTYLLDV	Branchiset Samle_seri	TUCKORRANTCOCATONIATILITEROBILITIAR ILCORCORTO AURIARIDA-TOCINALITIA INAPIDODENDOLI DAI INCOLUMIDITI
Gallus_gal Norodelphi		Macaca mil	GRAGAQOVVOLDEPT-QQPLHLVSPKELDVTIDAKLACOPRELIFLGEPKLACLARGLOP	Resopus 81	LKSBULQEQH1L	Spalax gol	YCRONOMACKYNKLIASTLAACEUTSENTNEHYLERALATKEDFLERONLDEFTLEN	Tytzaulog (exsetron)	SHURGH-KISINILISIYOAN LISEYOAN LISEYOAN AND MANYONYA YOOYA
Heteroceph		Homo_sapie	NEAGAGOVVULDETT-GEFLEURSISELEVILINGASDERELIILGIERLASIAGLE NEAGAGOVVULDETT-GGELEURSISELINTIDAKLASIERELIEURSEKLASIAGLE	Roolis car	LCCCMITCANTL.	3palax_car	YORVOOPHINTWILLINITAAGENTIINITMINYYIMINTATHEIFININTYILIN	Second A.	ADD.APTS-BECHDLARPTIEVILID-ADBCLARPVLINDVLINDVLINDVLINDVLINDVLINDVLINDVLIND
Cania port		Pongo_abel Somaarus 1	NACASSYNTEERIN - SOFULIYING IN THE AND AND THE AND	Tweniopygi Meleagris		Callithrin Macace mil	YEPOYOGNINATANGALEFILANGEYOGYTNIKYILINKTATKEDILINEVLICIFILIN YEPOYOGNINATANGALEFILANOOSITISIYTNIKYILINKTATEEJIILINOVLICIFILIN	Mulls our Textingypi	ALADERIN-HLIDTINARGIOTULIKI-KILITIDAARLUNINTIA ALADERIN-KILIDTINARGIOTULIKITKILARDIARDAGIOTU
Bos taurus Sus scrofa		Oryctolagu	GPROMEDAVELINEST-EXPROCUPATION FILTER AND PRESENT AND	Gallus gal		Nimo_sipie.	YEROVOQPRETAKKASELENCEPTERVTNERYTERETATKEDEERPOVEDEFERV	Weleegris Galles ool	ALALINYT-RUUPPLARENTPTLLP-RECEILINERGAGELUMPTER
Cricetalus	100	Londonta .	SPAGADD VELDPTT-DRPVKLVSPAFLBPTIDAKLASNEGPLTUSSSKAPLAAGLUP GPAGADD VELDPTT-DRPVKLVSPAFLBPTIDAKLASNEGPLTUTLOSPKLATLAAGLUP	Norodelph1 Heteroceph	LQLEARIQUVILL	Pongo_sbel	YERVAQUARTARCASICARGENTER/THEFTINERTARCSCARPACECITERY	Non-delph1	8283875-022962868762022-+68551386628622967873 8283675-02296286876792+-899622668286296585
Regressed	E	Eguns_caba Alluropoda	GPEGREOWELDFTT-ERPVELVSTRFLSLTEDWERTUPPELTELROPKLRTLAKELSP	Ceris porc	10.2007 get and	Nonascus_1 Orvetolami	YERWORKSHINGCASELKAGENIGON MINYLIKOFATIKESELKEVIGETER YERWORKSHIKKAGENIGETERGENIGENIGENIGEN	Detectorpt Carria more	AURUSTER - ELECTRONIC METAL/ARTICLAR ESPOSICIONETRATIONETRATIONETRA STATUMENTE - EL CLINIEL PREDVERTI LAR ECHOCOLISIENCE ACATELARIZZEN INVESTOR
Hattus_sor Spalax jud		Casis_laps	GPVQAED9RELDPUT-EXPVRQVSFAFLEFTIEASLATUPRFLTILLGEPKLATLASGLG2	Bis_taurus		Losodonta	YERONGRHEITYNYLTHUGAGONIDENTHEITLINHIATHEIFLNEWLETFILM	turia por Ito, tairus Jan mmla	ALALANTA-ALZUMALINARYDOVCDIA-DITALLIA INGLAGILANDIGTU ALALANTA-ALZUMALINARYDOVCDIA-DITALLIA INGLAGTUMARITU
Spalax gal			- ************************************	Bus scrofa Cricetalus		Equis daba Allaropoda	YEPOVOCPRORTVINEASELAAGOOVIDEVTNIKKYYLIKERIATVEDELISEVULDEESEV YEPOVOCPRORTVISILASELAAGOOVIDEVTNIKKYYLIKERIATVEDELISEVULDEESSY	Diportable	1000000-00000000000-000000000000000000
tpalax_gol		Branchiost	APLALOOTALOPLIPEPTED TTVLASOPLIO CALSHOLTSMEDLOPLASOTEHIOTY	Mus muscul	LOVENPPORTITI-	Caria_lupu	YOPOT SQPROM VELOCIT AND AND A LIFE ON THE CAME ACTUAL SAFETY AND TO THE SAFETY AND THE SAFETY A	Man Marriel Reffree mort	VULUETS-IL/INTEL/MOVETLLAD-FEDOLLETSLAD-FEDOLUSTS VULUETS-IL/INTEL/MOVETLLAD-FEDOLLETSLAD-FEDOLUST
Spalax car Callithrix		Danio_meri	MURENTROPUTATION OF AN AND A CONTRACTOR OF A CONTRACT OF A	Rattus nor	LOMENPPORILLL	0.0196.50	1.1 · · · · · · · · · · · · · · · · · ·	Nation per Spalae, jud	ALIELINITE-ALIELINGUNERVIENIUTE-LINICELININGLEOGRAMINIQUE
Recace mul		Tetraodon Creochromi	MTLRESGTRGEPHITATIKE	Spalas jud		Branchlort	TADARWANTAPOLY/ALGETESAWAGGAPELSDSPVATFMLDALOLALQLAU/AM90	ipalas pri	AUCENTE-RUCE QUINEMANDELINELONGLUNENQUINERAUGEU AUCENTE-RUCE QUINEMANDELINELONGLUNENQUINERAUGEU
Rono sapie		Renopus_E1	ATLANSTRUMULTOPIER	Spalas gol		Tacio reri	NEVLEW/EAVERORY/ALCETERA/GOOR/ALLOTY/ACPMIC.INLOLAR/LOLAR/ACLINIC/INC	Spalas par Califyring	ALIAL MATTER-KURLINGUNARY TATE UND LANGELLINTING AND ALIAN MODELLIN ALIAL MATTER-TLAL INSTITUTE (UND LANDAL LINTING AND TATE MODELLINE AND ALIAN
Pungn_abel Romascus 1		Roolis car Taesiopygi	GTURGETERUPERE	Spalas cur Callithrig		Tetraodon	GEVLEENIQUS PORPHILLOETISSA ISOGAASLISUTTI VAGHINLOKLISLAATLISLELVINIQ	Natara mil	ADCINITE-EXCHINISION/CAR-LONGLASSION/CARTINE
Cryctolage	NOVCEDNERLOLOCOPYPHENETRING CONNECTION CALONETRICHTER CALANCE.	Meleagris	GFLAPSOTECHLEPSTOC	Mataca mil		Creochroni Xenopus II	NEVYKYKLALECKSINLEETSIAJOGDAGLOOTYVAEHKLEKLGLGAALGLOVINGO QTIFQIVKESIPGKINLEETSIAJGGGEGLINTYLDEHKLEKLGLAAKSEDVINGO	hono hipin hongo kbel	ADMANTE-LAUPTERSKONDLARLAPELANDRANDTAND ADMANTE-LAUPTERSKONDLARLAPELANDRANDKANTERSTOP
Losodorta		Gallos gal	GTL8PG01970FL1P8P82	Romo_sapie		Anolis car	SEETCHOOL/ACCOUNTIEVED SYLDICAL STRUCTURE STRUCTURES	Sinsecus 1 Dryctalage	Advances - La Promision La P- La Robert galogram technologia Automotive - Robert Automotive - Descharting Automotive to
Eguus_cabs Ailurnpoda		Nonodelphi Neteroceph	ATLANSTRUMPERING	Fongo_abel	10128070202112	taeniopygi	YEVLETVOOTVPDKXWLGETSGAVDOGAPELINTVVADPMELIKLOLGANLODVVMBQ	Loundonty .	KINLWYTH-KIQUYTLIFUR/VIRTURPBOHOLLINDYLLADQCUM/SOFTU
Canis Jupu	<u>81</u> <u>82</u>	Caria_porc	ACLAPSICIAL FORMAL	Simascus 1 Oryitalagu	LALERY TOPOLLA	Welsagris Gallus gal	RONGE STEATY FORWARDE FEATOGRAPHILINTY A SPINEL DELGAAR CODY MIG RON LIGT VEATY FORWARDE SOAR SOCIAR GLAIRTY A SPINEL DELGAAR BOODY MIG	Bous table Alluranda	AUXLINETS-BLOCHTUPARIUMILTEBLHULLARING AUXLIANTORY AUXLINETTLALPSIL/200700000000
		Bos taurus	APLAPSONNEPPLIPPPORE	Londorta		Mocodelphi	QEVEQVIDETERIORVALGETESSARGOLAPELIOTYAADPHILUULOLAANHODEVVHIQ	Caniz_Tops	LUXURYTY-RULEPHOPSICPHOTECLOR-LOPORLUMINGLIGO/FURMINERTU
Ecanchicet :		Sus_scrofs Cricetalus	ATSPATTSPLIPPPE	Equis_caba Alluropoda	+LCERFFCELLL+++++++++++++++++++++++++++++++++	Saturneyk .	ORTOWERTH CONVECTOR AND A CONTRACTOR AND A		19 x 11 1 1 11 x 1 1 x 1 1 1 1 1 1 1 1 1
Denio_reri	<u>83</u>	Mus Muscul	ANDARGOTATIOFLIPOPORE	Canis lupu	LOCARETORIUL LCCERTORIUL	Cavia_porc Bos taurus	QRILQVIETH-ORVILLETSSATOGRAFLLETTAGPHLDLOLDAQHUDVVHQ QRTLRTHEXINFLKKVLLETSSAFOGRAFLLINTFAACPHLDLOLDAQHUDVVHQ	Beauchinet.	statement were and a second se
Tetraodon Greochromi		Battus nor	ALTRACTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTU	0.000		Sus_scrota	QUE PONTRETT PLACEMENTE DAYS COLAPSEL INTERTIMENT AL SUBJECT OF AN AND A DAY OF A DA	Danie_reci Detración	LEGEK, MAREK, QL MERELARY (PTORQAND: 7 ++)
Netopus_01		Spalar jud Smalar mal	ATLANOTYTIST. FORME	Branchilest	WEGLEFEAMETERS/FRAFTMER/EMERIENER/COLDF1FELW/LINESL-NEPERA	Cricetulas	privativeste potential cessor and a substance of the second	Dreichnei Mongus 31	UKALLINAKU/ANTUNYUNGANCO UKALINATU/NUNUNNINAKALKI
Anhlis mar	***	Spalas_pal Spalas_pol	ATURPOORTUPUTONAL ATURPOORTUPUTONAL	Danio reri	RESULTERNET INFORMATION CONTRACTOR AND A	Mus muscul Rattus nor	QELINYHETTYKKYALKESSAYLAKAPLLINYTAADMILKILAADMILKAP	Really sur-	LTHMPLENGHLELPHERYCPHYCHNAPPINCT++
Taeniopygi Helesgris		dpalan car	APLAPOOTKTIPLIPOPKE	Tetrandon	NEELQKY1N/XXTETYVQLAATNICHLULVYLUALLITADINNIDISA	Spalax_jud	OKILOVVERTRPORYVKLOETSSAYOOUAPLLINTENADINMLUKLOLAACHSCEVVHDQ	Testingys: Releastis	DEFICIPATION CONTRACT-
Callus gal		Callithria Nacaca mul	AFLARGITETETLIFURGE	Creschroni Netopus 31		Spalaz gal	BLITGAREELINGIKARTEELINYSSOODITTIILENYSWITTIKTOTTIVGHUTEAAMAG	Gallar gal Monskipli	LIGHCARDTUCIARDOP/CRAMUNCT-
Minodelphi	10 m	Nono saple	1/130/07/170/11/19092	Anolis car	RECOMMENDETTITESTVOILUMERACACLALIFCINALLARCESCHORTHA	Spalax_gol Spalax car	QELIQVEETROXVELETISA/GRAFLISETRAGHELDELGISS/RETEVING QELIQVEETROXVELETISA/GRAFLISETRAGHELDELGISS/RETEVING	Belarcoph.	FUNCTIONALISAAANAN CAACAMA CONTRACT
Heteroceph Cavia porr	111	Pumpo_abel	AUXPOTRT9/119900	Taenitpygi		Callithrig	ONTOWESTING ON MICETELAY COCASELED AT A COMPLEX COLORISON OF	Canla_port hts tarra	LINE HARDLED AFY SPECIES IN ACCEPT LINE HARDLED AFT STRATEGY (MARINE)
Bos taurus	1	Somescar_1 Septimizes	ATLAPATTURU POPUL	Meleagris Gallus gal	RESERVISIONTI TENTLO LLETERSCOGPEL VELLANAGLARAGORISIA 	Macaca_mil	ON NOVESTRAGONILISTISA YOGA PLISTRAGING DELO SAMESEVING ON POWERTRAGONILISTISA YOGA PLISTRAGING DELO SAMESEVING	tus errota	LINE CONTRACTOR CONTRACTOR
Sus scenta		Losodonta	APLAPSONTOP_IPOPUL	Monodelphi		Homo sapie Pongo_sbel	OULOVERABORINTELEPERACENTTELEVERY TROUBLE TRANSPORTENSES OULOVERABORINTELEPERACENTTELEVERY	Crostalus Mux manul	LIGHTLAFEDLAMATENTY/INHIGHL
Cricstilos Mas mascal	***************************************	Dezus_caba	ATLAPOOTATISFLIFOPHE	Netszoceph	REQTORENTTTYSREEVORLISERACEGURLIPGLEELLATACOORCESSA	Nonascus_1	ON FORVESTIFICATION LOEDSSATIONAPELLIOTENA/EMILINLICLIA/HICEVINIO	Ratting nor-	LINF, HULLANDY OV INALIACI-
Rattus nor		Ailuropoda	ATLANSTRUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTU	Caria_porc Bon taurus	REDVOK-PRITTY BROWNELI BRACECULT FOLKELLATADONOLOSIA BROVOKY FINITY BROWNELI TRACECULT FOUNDALLATTONICOLOSIA	Oryctolage Louissers	protocontensional constantical international dependences	the market	and a second state of the second state of the second second second second second second second second second s
Speles_jud	Rhemmen	Canis_lopo		Dis scrofa	NOROFONTTIREVOLUTIRECELULIFELIALLATEOUNDUSIA	Equis_cabe	ONYOWETTERMINELETSSA STOCAPILISES PARTINE DELCEMENTERVINO		
Ipalax_pal Spalax_pol	82			Cricetalus		Alluropoda	Condivation Report to 2010 Text Arrange Determinants	Spalax_gal	
ipalan car		Branchiost Danio reri	RETHERVERORIASPTHCVGYTHTTFGAVLLANGSLEDPENAGLLANTAAAGPKVMEL EKCESAPQOI	Mes_muncul Rattus nor	 ROTORE PROTECTIONS VALUES ACCOLLETED ALLATED LANKS DIA PROVINE PROTECTION OF ACCOLLETED ALLATED LANKS DIA 	Canis_lips	0030	Spalax_gol	
Callithria		Tetracion	ELTFERINGS-	ipalas jud	ROTTREFENDTY FROM INCUT FRACIDLELIFSTRALLATADERMICERA			Spalax_car Callithrix	
Macana sul Homo sapie		Greochromi	NCEANES	Spalaz gal	Reproversion of the second structure of the s	Branchiost	ELFEA-WALLINGLIPLJDIWLELLINGLVGENVLKNVTDGEDWVERBE	Macaca mul	
none repre		Xesopum 31 Analis car	LTTBQATE	Spalas gol Spalas car		Danio reci Tetrandon	VLDAFTHE VEHILDER	Homo sapie	
200 million - 10 million - 10				1000		11110000		Pongo abel	
Putpp_abel		Taeniopysi	TURENAGE	Callithria	······································	Greochroni	YL0000195/Y004L04L9209L04L095/W0E94E922Y15E	Nonascus 1	
Tomascus_1 Cryctolage	THEFT WARD SCHOVEN TO CONCERN FRAMEWORK AND CONCERNED AND A CO	Meleagris	11.21.3.4D	Mácaca mil	REFERENTIATION INTRACTION PROVIDED AND AND AND AND AND AND AND AND AND AN	Menopus 51	ALFCASTINE VELIFIER	Oryctolagu	
Losodonta	82	Gallus gal	VL3EF34+R2	Hono_saple	NERTQUEFNEETTERSEVOVLITERSCOLLELERALERTACLORUSERA	Atolis car	VLRGAGINELVIAGPERLF	loxodonta	
Eprus_caba	7U	Monodelphi Netaroosph	10010000- 50050000-	Pongo_abel Nomascia 1	EFFY OR FEND THREE VOLUTERS COLOL I FOLKALLETAD LINE ON EFFY OR FEND THREE VOLUTERS COLOL I FOLKALLETAD LINE ON	Taeniopygi Nelsagris	VERGADTINE VERSTERIE VERGADTINE VERSTERIE VERGADTINE VERSTERIE VERGADTINE VERSTERIE	Equis caba	
Allutopoda Canis lupu	Ref 112	Carla pero	\$3859702	Oryctolagu	REDORETROTTIKSTOMAAFARCISCHITFICHALLETAOLOMISINA	Gallus gal	VERGALITIKENTREFTELT	Alluropoda	
		Bos_tairus Sos scrofs	EDUCATION CONTRACTOR CONTRAC	Losodorta Equis caba		Munodelphi Heterooeph	VERGEBURGERUNGEN VLF	Canis_lopu	
Inanchiost.		Cricetalus	1905/982	Ailuropoda	NOROPOLINETTEREPUBLIAFIACISCULTFELIALLITACLONIUM	Cavia poer	VERONOMIEUTOREWERDIMENTEROLWITTWINNINTOR		
Dahlo peri		Mis muscul	\$78522982	Catis lupu		ftos taurus	VLPGAGHTEL/DGHTEPLF		
Tetraodon		Rattus nor Spalas jud	579(5)(50) 579(5)(50)		· · · · · · · · · · · · · · ·	Sus_acrofa Cricetalus	VUPGACTINUTORITPUP		
Cosochrimi Nancpus Si	ALLICALCONTY	Spalar gal	59032000	Beanchipet	QULIANTAARDPROMELONEPHILORI SIRT MATTLULOPHOLISUUSEAPHNEEL	Mus moscul	WYCADNE PERFERT		
Anolis car		ipalas_gol	200020000	Danio_reri	#111xycessgowerweitoxepatrexxxovantorgotoxemailxgiligestitweitot	Battus_ter	VITCAGNYCUCENTER LP		
Twenlopppi		fpalaz car Callithria	STROUVIED	Tetraodon Desocheomi	ASULASTRAASTRADELOREPROTEXXXILALIGADUGEDETVLAKILASIAFYROAD. ALLOGYTYDAHTYNEWELOREPROFEXXXIIX/DYYQUDLOPIALIASMREEXLYQOAD.	fpalax_jud fpalas pal	VFFGGSYELVERFEFLF		
Heleogris Gallus dal	**************************************	Macaca_mil	100000000	Xecopus_31	KLLIDHCILAKYNCAMELONEPHOPPOKOSTYLISSKYLOKOPHULHULDHOPPIKOSIL	Spalax pol	WYGARNELISBUTHLFDINLELLINELVERVERWER		
Monodelphi	**************************************	North saple	219002/100- 519002/100-	Acolls_cur Taeniopygi	QLI LOTTO QUE INDRELO REPORTANTO EN LOTTO O QUO DE LA QUENTE EN ENTRE. QALLOTTA QUENTI ENELTRE PREPORTANTO EN LOTTO LO QUE O QUENTI AL VOLTO L	Spalax car Callithria	VYTGARNELVERFELT		
Heteroceph Cavia port		Pungo_abel Nomancus_1	518(20/80)	Taeniopygi Heleagris	AGTTELEVERAL METERETERENE AND ALT DE LEVELET AND A	Callithris Mataca_mil	VFFGAGETEL/CERTPLI		
los tairus		Orottolagu	519(3)(79)(2	Gallus gal	RELETERARIES SELENE POPERARIES CONFIGURATION AND REPORT	Nomo_sapie	VTFGAGNELITERTPUP2016.01LPX1VCT010484VQ68		
bus_scrofs		Lexodonta Egun caba	519(55)/9(5)	Monodelphi Heteroceph	FLPLOYTOFFNMSKELCKEPSTRKSTIVIDSSCUPPINIKILIFYD-1-KAVL SCILLOYDSKEPHI SWELCKEPSTRKSKEIFIDSLSCUPPINIKILSEST-PROXI	Pompp_abel Nomancus 1	VPFGADIVELIDDRPDH2DVRLELLIPRELVDRVVDRSVRD VPFGADIVELIDDRPDH2		
Cricetulus Mus mostul	111	Ailuropoda	200/20/10/2	Cevia porc	ollowisekorotskeldkelekerskoodtribokolokorolakilokiit-Ikkaki	Gryctolagu	X1/5A01021/08/07/02		
Rattus_nor		Catis_lopo	19000400	Bes taurus	QL222055508YM158E208E208F28H351F1958QL5E20F1E78H1L0858-F9888L 0111 https://www.tiset.com/sectors/internet/in	Londont.a.	VITGADINUTDERPLI		
Spalax jud Spalax dal	ACCERCASSIGNORPORTALSULLREREPLAYOTIA			Bun_scrota Cricetalus	QLLDVCASNYNISKELONE NOFWORKIY JORYGAEDY JOLHLDWSA - PROAKS QLLDYCASNYNISKELONE PHOPROXACT JOCQUEED Y OPHILDWSA - PROAKS	figuus_cabe Alluzopoda	VTTGARTELVORTEL		
Spalan_gal Spalan_gol	REFERENCESSIONCENDERLELL	Branchinet	OVCREDEHORSTHOROUTELLINGHAMOREFULATOHICONTL	Nus_muscul	OLLEW SEARCHISSELCHEN SPARKATLINGSOLGEN VELHOLLOKSA-FONARL	Catia_lopu			
Spalar car		Danio_peri Tetraodon	-572.223	Rattus_tot Spalas jud	QUILENTSONDHISHELGREPHERMAGISSIGLQUEDPVELMHLIGKSA-FORMU. QUILENTSONNYSIBHELGREPHERMAGISISISIGLQUEDYTELBALLOKOT-LANNKI		***********		
Callithria Entern mil	LINEPLINER	Tetraodon Oreochroni		Spalas_you	QUITERCONDUCTIONED METADORY OF A DOT VETADOR OF A DOMAIN	Branchinet	CONTROLOGICAL INVALIDATION PO		
Habaca mai Rumu_sapie	TTLLL	Xenopus_51	-100	Spalax gol	QLLIAPCEBRATISTEMES/GREPHOPMODALISIDGEGLGEOTTELHILLIARIT-LADVIS	Danio_reri	SQLOOPINYLICTNOSIC		
Pongo_abel		Anolis car Taeniopygi	YCR	fpalaz car Callithriz	QUILINTEENNYEI INELENENETNENEN TETEOLOGISEENTELNET-LINNE GUILINTEENNYEI NELENENETLENETLENETTIKSSSSSERTELNET-MUSAEL	Tetzaodor Greochromi	PAGEENALTUKONKORCILLOFUNKONEAAPSUVILCIINSOAATUK FURININYI KUNIK		
Rohandus 1	CHICKNER CONTRACTOR CONTRACT	Taeniopygi Meleagris		Hacaca mul	CLILDYCORPORELORING FLORE NO FLORE CONTRACTOR STATES AND	Xenopus 31			
Cryctolagu Losodosta	QUARCANCELITIPOLIPOLIPOLIPOLI-CO-CO-CO-CO-CO-CO-CO-CO-CO-CO-CO-CO-CO-	Gallus_gal	VE	Nono_sapie	gLLL2xtstantratiset.detroiruxxxtrtacigLatorigLantLanst-moart	Anolis car			
Equip caba		Monodelphi Weterooeph	TC	Pongo_abel	QLILDICISHIYINI SHELOHEPISIYLIKAD LIYINGOQLIELIYIQCHILLIKIST-FIXAAL QLILDICISHIYINI SHELOHEPISIYLIKAD LIYINGOQLIELIYIQCHILLIKIST-FIXAAL	Tanniopygi			
Ailurspeda Canis lups		Ceria goro	7CE	Nemascus 1 Oryctolagu	Qualification of the control of t	Melsagris Gallus gal			
carra subs	10/Larrie ALA	Bos taurus	**102**********************************	Losodoeta	KLILOKOBKSYKISHELOHEMBFLAKKETYIDOFQLAEDFTELHELLOKSH-FKAND	Monodelphi	+EIXESCULT		
123302542		Sus scenta	-VCE	Epus cabs	QULLOCCERTRIBUTIONE PRIPARAGENE DELOCATION CONTRACTOR CONTRACTOR CLUCOCCERTRICE SECTOR PRODUCTION CONTRACTOR CONTRACTO	Beteroceph Cavia more			
Brinchinst Datio reri	++ASQUARTVEWE+CARMADISATISATISATISATISATISATISATISATISATISAT	Mis Miscil		Catis lupu	ALLEW SHARE SELVE SHE WITH WATEFING AND A SHE	davis porc Bus taurus			
Setracion		Rattus bor	105		11 * . I I I I * I	dus scrofs			
Creochromi	GEPOROPOLILINOV-GRADRASTOPHYLLINYTIDAGLASS/ESTVALLISROOPTLAAALISP	Spalaz jud		Branchiast	ACPOTTNYONTLAFLOR	Cricetalus Nos muscel			
Banopus SL Banlis can	GRATPELTIDLATHI-KSIINFRELIVTVDGLADERVIAR-ROOMTLARGLIP	fpalan gal fpalan gol	TCB	Denio reri	NEPOTRAVERSCLATOFICS INCOMPANY INCOMPANY INCOMPANY INCOMPANY INCOMPANY	Mus_muscul Battus cor			
Acclis_car Tecnicoyol	EAGELLERADEALAQUEETUALTIANCADIENTATUDACUETUALUE TENDETUNALULTUALUAUNULUUUUUUUUUUUUUUUUUUUUUUU	Spalaz_car	101	Tetraodom	FIPOVOZPROREIOTUROFUQIOARAVOACTINETYLDIREALUDOFUDIOVUDTURERI	Spalar_jud	BERTERCHINFBERGELLE		
Meleagris		Callithria	-108	Greachrani	TEROVOPHORIVOLLOGIQUCAMAIDACTREETIVAGADITEROFLORINIDELALKI	dpalax gal			
Gallus_gal		Macaca mil Homo sapia	+1CE	Nemopus_S1 Realis_car	PUPOLOGPHYLOGPHEASTLEYWGE INS TWHYTYWOUTABEED'I SAOLLOTLANGE I YSTUMOPHNYTYRLEASTLEYWGY I DYMYNHYTLAGATATHEDHNS FYNLOTFI I AW	Spalax_gol Spalax_car	BC-SCATTLACTSINGBCGSSGTLT		
Nosodelphi Heteroleph	VIERLEVIKARE-EVIVENERFALATTIDAKARENITITAKERE NAKOEVILIKE-ENIVENELEVITAKERITIDAKERITERENITI	Fongo_abel	+-10X	Taeniopygi	YSPOVOGPRINTORLASPLASOGRY LOSVTNIHT I VINGSATROOPLSHEVLDTPATAT	Callithrix			
Cavia port	BBALREVERNPLAPTIONICATIONPELAPLICENCIATI	Nonascus 1		Nelsagris	YSPONOQVARATOPILLASPLASOGIALOSVTNEHYYNGRISATREDFLSPONLOSPATAN	Macaca_mul	ALTERNAL AND A STREET, I S		
bus teurus	AAA/NJGAAELEFTT-ERPLACESPARLSPTIDAKLATSPRIPTTYJJSSRUATLARELAP	Oryctolagu Leendorta	+ [12	Gallus_gal Mocodelphi	TULDVOQPROHTORLIASTRASOGRAIDSYTNEHYTYNGASATREDFLSHEVLDEFRAAT ROPOLSOSRITSGELLASPLEVUGOZYTORYTNEHYTYLMGROATREDFLSHEVLDEFTFEA	Nomo_sapie Nompo_shel			
Sus_scrofs Cricetalus	REAR NOVINGENT - GRIVELISEATINT DORATORIELTS INFORMATIANSIST STAARDOVINGENT - DRUK DEPENT INTERNETTING ATTORNET TO STREAT LANSIST	Epus caba		Heteroceph	REPORTORING STOREN DE TREKTIVER LARDELE PORTORI TAL	Sonancis 1	······································		
0.000.00000		Ailuropoda	108	Canta_pore	HEROTRÖHHORRANDORLÖGELÖGENDORLUNGELINHEN LINDELINGEN INTERDALINHEN LIDTET MAN	Oryctolagu	Be-GERYLICTRING		
r •			10.1				(.1		[01 1]

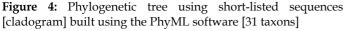
Figure 3:

MUSCLE (3.7) multiple sequence alignment [CLUSTAL format] of short-listed sequences

[31 taxons]

BLAST algorithm was used for comparing primary biological sequence information; this preliminary fundamental program uses a heuristic method to find homologous sequences by locating shot matches by seeding followed by local alignment. The main idea of initially using BLAST is its ability to produce high scoring sequence alignment between the query sequence and database sequences. The BLAST algorithm gives a very good speed and relatively good accuracy. However, it has been widely believed that BLAST should be used as a first pass sequence alignment. Phylogenetic trees constructed in this study indicated that HPSE protein is conserved and may play and important role in organismal evolution (Figures 3 & 4). It is interesting to note that the conserved regions as shown in Homo sapiens are similar to those found in some other organisms that have this conserved gene. Presence of heparanase in the major organisms indicates that it is crucial for the development of the physiology of cell membranes. Its high conservation at certain domains indicates that its function is preserved. This gene was absent in lower organisms like the invertebrates and the prokaryotes. In conclusion, the evolutionary relationship of HPSE gene was established based on the sequence alignment, conserved sequences and phylogenetic trees. The results of the published data on protein sequences of the above genes showed that the sequences are highly conserved especially at certain domains. Human sequences consistently clustered with their mammal orthologs within these genes clearly indicate the importance of these genes in evolution.





Acknowledgement:

Authors would like to extend their appreciation to the Research Centre, College of Applied Medical Sciences and Deanship of Scientific Research at King Saud University for funding this research.

References:

- [1] Sanderson RD *et al. Matrix Biol.* 2004 **23**: 341 [PMID: 15533755]
- [2] Vlodavsky I & Friedmann Y, J Clin Invest. 2001 108: 341 [PMID: 11489924]
- [3] Abboud-Jarrous G et al. J Biol Chem. 2005 **280**: 13568 [PMID: 15659389]
- [4] Dong J et al. Gene. 2000 253: 171[PMID: 10940554]
- [5] Fairbanks MB et al. J Biol Chem. 1999 274: 29587 [PMID: 10514423]
- [6] Cohen E et al. FEBS Lett. 2005 579: 2334 [PMID: 15848168]
- [7] Dempsey LA et al. Trends Biochem Sci. 2000 25: 349 [PMID: 10916150]
- [8] Nadav L et al. J Cell Sci. 2002 115: 2179 [PMID: 11973358]
- [9] Friedmann Y et al. Am J Pathol. 2000 157: 1167 [PMID: 11021821]
- [10] Benson DA et al . Nucleic Acids Res. 2011 39: D32 [PMID: 21071399]
- [11] Thompson JD et al. Nucleic Acids Res. 1994 22: 4673 [PMID: 7984417]
- [12] Saitou N & Nei M, Mol Biol Evol. 1987 4: 406 [PMID: 3447015]
- [13] Altschul SF et al. Nucleic Acids Res. 1997 25: 3389 [PMID: 9254694]
- [14] Grishin NV, J Mol Evol. 1995 41: 675[PMID: 18345592]
- [15] Dereeper A et al. Nucleic Acids Res. 2008 36: W465 [PMID: 18424797]
- [16] Edgar RC, Nucleic Acids Res. 2004 32: 1792 [PMID: 15034147]
- [17] Castresana J, Mol Biol Evol. 2000 17: 540 [PMID: 10742046]
- [18] Guindon S & Gascuel O, *Biol.* 2003 52: 696 [PMID: 14530136]
- [19] Anisimova M & Gascuel O, Syst Biol. 2006 55: 539 [PMID: 16785212]
- [20] Dereeper A *et al.* BMC Evol Biol. 2010 12: 10[PMID: 20067610]
- [21] Chevenet F et al. BMC Bioinformatics. 2006 10: 439 [PMID: 17032440]
- [22] Marchler-Bauer A *et al.* Nucleic Acids Res. 2011 39: D225[PMID: 21109532]
- [23] Parish CR et al. Biochim Biophys Acta. 2001 1471: M99 [PMID: 11250066]
- [24] Wu CH et al. Nucleic Acids Res. 2006 34: D187 [PMID: 16381842]
- [25] Strait DS & Grine FE, J Hum Evol. 2004 47: 399 [PMID: 15566946]
- [26] Jones DT et al. Comput Appl Biosci. 1992 8: 275 [PMID: 1633570]

Edited by P Kangueane

Citation: Shaik et al. Bioinformation 8(9): 415-419 (2012)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited

Supplementary material:

Table 1: Accession details of short-listed sequences

Accession	Description
<u>NP_006656.2</u>	heparanase isoform 1 preproprotein [Homo sapiens]
XP_002814980.1	Predicted: heparanase-like [Pongo abelii]
XP_003265886.1	Predicted: heparanase isoform 1 [Nomascus leucogenys]
XP_001104975.1	Predicted: heparanase isoform 1 [Macaca mulatta]
XP_002745691.1	Predicted: heparanase isoform 1 [Callithrix jacchus]
XP_001493332.1	Predicted: heparanase isoform 1 [Equus caballus]
XP_002912532.1	Predicted: heparanase-like [Ailuropoda melanoleuca]
XP_003414156.1	Predicted: heparanase isoform 1 [Loxodonta africana]
CAJ30017.1	heparanase [Spalax galili]
CAJ30018.1	heparanase [Spalax golani]
CAJ30019.1	heparanase [Spalax carmeli]
CAJ30020.1	heparanase [Spalax judaei]
XP_003511884.1	Predicted: heparanase [Cricetulus griseus]
XP_002717028.1	Predicted: heparanase [Oryctolagus cuniculus]
NP_001139602.2	heparanase [Sus scrofa]
NP_776507.1	heparanase precursor [Bos taurus]
EHB00176.1	Heparanase [Heterocephalus glaber]
<u>AAN41636.1</u>	heparanase [Mus musculus]
NP_072127.1	Heparanase precurssor [Rattus norvegicus]
XP_003469452.1	Predicted: heparanase-like [Cavia porcellus]
XP_001376885.2	Predicted: heparanase [Monodelphis domestica]
XP_003221281.1	Predicted: heparanase-like [Anolis carolinensis]
XP_002190948.1	Predicted: heparanase [Taeniopygia guttata]
XP_003205640.1	Predicted: heparanase-like [Meleagris gallopavo]
NP_989498.1	heparanase precursor [Gallus gallus]
NP_001120267.1	heparanase [Xenopus (Silurana) tropicalis]
NP_001038470.1	heparanase [Danio rerio]
XP_003434063.1	Predicted: heparanase isoform 1 [Canis lupus familiaris]
CAF94326.1	unnamed protein product [Tetraodon nigroviridis]
XP_003453646.1	Predicted: heparanase-like [Oreochromis niloticus]
<u>XP_002603693.1</u>	hypothetical protein Brafldraft_93100 [Branchiostoma floridae]